# PATENT ABSTRACTS OF JAPAN

(11)Publication number:

01-220063

(43)Date of publication of application: 01.09.1989

(51)Int.CI.

GO6F 15/20 GO6F 15/38

(21)Application number: 63-046509 (22)Date of filing:

29.02.1988

(71)Applicant: FUJITSU LTD

(72)Inventor: TAKAHASHI YOSHIFUMI

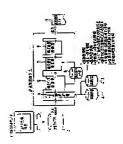
# (54) SYSTEM FOR EVALUATING DOCUMENT QUALITY

# (57)Abstract:

PURPOSE: To quantitatively evaluate a quality in the syntax and meaning of a document with the use of a computer system by calculating the quality in the syntax and meaning regarding the inputted document as a

weighted grade.

CONSTITUTION: A Japanese language analysis part 4 to analyze the inputted document and to form syntax information necessary to evaluate the quality in the syntax of the document such as predicate number information, compound sentence and complex sentence information, double negative information, etc. and a syntax grade calculation part 7 to obtain the weighted grade for the quality in the syntax of the document based on the syntax information necessary to evaluate the quality in these formed syntaxes are provided. The total sum of the grade calculated by the syntax grade calculation part 7 and simultaneously a guide-line sentence to correspond to the evaluation in accordance with a necessity are written. Thus, the quality in the syntax of the document and the quality in the meaning can be quantitatively evaluated.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of

rejection

[Kind of final disposal of application other than the examiner's decision of rejection or

application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's

19 日本国特許庁(JP)

① 特許出願公開

## 平1-220063 ◎公開特許公報(A)

இint Cl. 1 G 06 F

庁内整理番号 法别記号

@公開 平成1年(1989)9月1日

15/20 15/38

V-7165-5B X-7313-5B 302

審査請求 未請求 請求項の数 2 (全11頁)

文書品質評価方式 の発明の名称

②特 順 昭63-46509

**勿出 顕 昭63(1988)2月29日** 

神奈川県川崎市中原区上小田中1015番地 富士通株式会社

神奈川県川崎市中原区上小田中1015番地 畜士通株式会社

弁理士 井桁 貞一 外2名

1. 税明の名称 **女童品管牌值方式** 

## 2. 特作環状の範囲

(i) 文啓の株文・志味上の品質を評価する文書 品質評価方式において、

入力された文書を解析し、透露数情報、建文・ 複文情報、二重否定情報などの文書の積文上の品 質を評価するために必要な構文情報を生成する日 本語解析節例と、

これら生収された構文上の品質を評価するため に必要な権文情報に基づいて、文書の構文上の品 質に対して加重つるの評点を求める領文辞点算出 前のとを伝え、

この様文評点算出部のによって算出された評点 の途和および必要に応じて評価に対応する指針文 を併せて記載するように構成したことを特徴とす る文書品質評価方式。

23 文容の様文・意味上の品質を評価する文書 品質評価方式において、

入力された文泰を解析し、不被一な文体情報、 **冗長要項情報などの女に関する実味情報、および** 用語の扱り抜雑、不適切な用語情報などの用語に 望する意味情報を生成する日本語解析部200と、

これら生成された意味上の品質を評価するため に必要な意味情報に基づいて、文書の意味上の品 質に対して加重つきの評点を求める意味評点复出 財団とを構え、

この意味神点算出部のによって質出された神点 の絶別および必要に応じて評価に対応する難針文 を併せて記載するように構成したことを特徴とす る文書品質評価方式。

## 3. 発明の詳細な説明

#### (模型)

文書の構文・意味上の品質を評価する文書品質 経価方式に関し、

文字の親文領報および意味情報に対して加速つ

# 特別平1-220063 (2)

さの評点を耳出し、これら深点に基づいて損文上 および重味情報の品質を定置的に評価することを

人力された文章を質折して博文情俗および意味 情報を生成する日本越新的版型と、これら生成され た様文上および意味上の展文を採用するために必 受以協文情報および意味情報に基づいて、文章の 様文および意味上の品質に対して加度つちの辞点 を求める様文評及正出版および意味再点互出版と を図え、この确文評互至出版および意味再点互出版 形によって当出された評点の疑知および必要に応 して評価に対象する精神文を簡単で記載するよう

## (産業上の利用分野)

本発明は、文書の様文・掌張上の品質を評価する文書品質評価方式に関するものである。

(従来の技術と発明が解決しようとする問題点) 文書、修に技術文書の構文上の品質および意味

1 (は、人力された文音(技術文書など)について、日本保護者(-1かよび日本経文法 - 2を 参照して解析し、文書の間次上の品質を指する ために必要な議員監督様、資文・限文管領などの 株文情報、および文書のを集上の品質を経済する からに必要な不詳一な文体情報、用語の割り賃権 などの宣都要な不詳一な文体情報、用語の割り賃権 などの宣都環幕を進奪するものである。

構文財疫算出部 Tは、人力された文書について 耐えば文単位に、波道改情報、黄文・複文情様な どの伟文情報に対して、加度つきの許点(指数 個】を実→質出するものである。

意味料点車形器11 で、入力された文書について例えば文単位に、不該一位女体情報、元長東東 情報などの文に関する意味領域、力よび用取の思 り情報、不遵切な用語報などの用語に関する意 味情報に対して、加減つ8の計点(物数像)を表 ・室出するものである。

#### (作用)

本発明は、第1回および第1回に示すように、

上の品質を評価する手佐が望まれている。

これら博文上の品質なよび意味上の品質の評価 は、従来、人手に切っていたため、下記のような 問題点がある。

(i) 評価の特度を高めために、評価する項目を 製工(設定すると、人手による評価作業が高大と なってしまう。

は 人為的なミスによる親れが発生し、この過れに対する防止策が必要となる。

は 評価が主観的であり、他の文書または他の 評価者の評価結果との比較が困難である。

本発明は、文書の様文情報および意味情報に対 して加重つきの評点を選出し、これら呼点に基づ いて様文上および宣味情報の品質を重量的に評価 することを目的としている。

#### (問題点を解決するための手段)

第1個および第1回を参照して問題点を解決す るための手段を説明する。

第1回および第7回において、日本語解析師 4、

日本道路所間4.14が入力された文理を繋折して 可論文解却よび意味情報を生成し、限文資品室 出版でおよび意味用真互出版を17からし進さ たた相文情報がよび意味開促を可いて、様でよ の加度つちの材点がよび意味上の加重つちの呼点 そ至出して出力すると共に、必要に応じても昇低 に対応した指針文を付加して出力するようにして

促って、人力された文書件に技術文書について 様文上および意味上の品質を加変できの評点とし て江出することにより、計算機システムを用いて 文書の例文上の品質かよび意味上の品質を定便的 に対価することが可能となる。

#### (実験例)

まず、末1回から第6回を用いて、本発明に係 わる文書の様文上の品質の評価について評価に従 明する。

第1回において、日本文作床デバイスしは、日 本語リードプロセッサなどであって、日本文によ る文書を作成するものである。

文書ライブラリ2は、日本文作成デベイスしを 用いて作成 (日本追販務数据) した日本文を保管 ・登録するものである。

請文データ収数部5は、日本感解所額4によっ で日本無限額4-1、日本額支洗4-2を参照し で解析された結果のうち、文章の請文上の必要を 評価するために必要な請文派等データもその変す ものである。この確文解析アータをは、因示の ように、当直数性何などがある。

博文品質評価部をは、博文界点変出等!によって、博文解析データをに基づいて重みづけして立 地された解点の値句を求めて出力すると表に、こ れら多々の得点(第3回から第5回を用いて扱道 する)に対応する指針文を出力するものである。

品質解碼リスト9は、様文品質評価値目によって質出された評価点を、例えば第6間右端の評価 を 32°として示すように、1文単位にリスト として印刷したものである。曲、この際、評価点

味している。

図中のは、天即解析、様文解析、様文領権の収 気を行う。これは、文節および様文の取得を行い、 运移鉄、重文、直文、建成本文、受動脈、進用中 止体、原政取開、二重で定文、"ように"のを定、 重通・自動機の需像、構文振りなどの様文情報を 収取するととを単している。

図中のは、多々の評価項目ごとに対点を計算する。これは、"女の間文"における情報項目の計 関端果に対して、多々の評価項目の許可範別を考 感し、後端する罪を関に示すように対点す。など を計算することを重要している。

図中のは、文の正しさのば点を行う。これは、 図中ので計算した多々の評価項目ごとの評点に対 し、多々の型みづけした後に、これらの値を総計 した評価生を算出することを重楽している。

図中のは、運出結果を協議し、帳票を作成する。 これは、時られた結果を関係し、品質体係リスト 9を印解、例えば第6回に示すように印刷することを意味している。据6回は、文単位の様果例を

といせて、複針文を戴6回中央に示すようにメッ セージとして印刷する。

次に、第2回フローチャートに記述した順序に 姓い、乗3回から第6回を用いて、第1回構成の 動作を詳細に観明する。

第1回において、図中のは、文章の入力を行う ことを示す。これは、第1回の文章ライブラリ 2 から彼み出した文書等に技術文章を、品質評価ツ ール3を構成する日本品解析版4に入力する。

関中のは、文書要素の解析を行う。

図中学は、大きの選々の要素の慣程を収定する。 図中学は、大の切り出しを行う。これは、大き (技術大きなど)の中から、一文づつ切り出すことを意味している。

図中のは、単語を抽出し、単語質和の収集を行う。これは、図中ので切り出された太から単語を 抽出し、更にお好をはがし、・抽金語・・ 性様 用語・、・ 鳴路・ などを計画すると共に、・ ワ ら がな・、、カタカナ・・ 、 頂本・、 デキ・、 ・ 後春・・・ 神枝起导・などを計画することを意

以上の処理によって、人力された文容の日本頃 解析が行われ、文の権文解析データをに基づいて、 文書の相文上の品質の好価人を声出することによ り、文書の相文上の品質を定置的かつ書版的に評 版することが可能となる。

第3回は、評価高級例を示す。回中機輸は用版 版(X、)を受し、凝棒は肝皮(Y・)を支す。 示3回(4)に対いて、固中のは片字原理であ で満点を与える結晶を走し、図中のおよび図中 のは関示直域によって変されるような解点された 肝点を与える短面を変す。この図中のの計算系 は、予め手限の文質(対人は、技術であり は、予め手限の文質(対人は、技術であり 通知であって、指文上の当様を出す。対して 通りであって、指文上の当様には対して では、大変には、形成となるよれも応うとは 成となるない。の図のからない。 のたるに、のように表して、図中のの か容異調の質測に、明点および上環1、20個中の か容異調の質測に、明点および上環1、20個中の か容異調の質測に、明点および上環1、20個中の か容異調の質測となるように異様ではんだ、回 もガグラフを評価基準として予め作成したものである。具体的に言まば、図中にデオメンに、評価 項目・高温数・に対し、1つの文中に記載されて れる活は数が入ばしないして3回の文に対して対 点を与えもよいと多数の文章を登録して対別した 場合、図示のように「・・し、」・3となり、 両者を組んだ直接が即かる即等項因となる。そして が成の会のよびのの機能観点なる。

また、第3回(の)は、第3回(イ)回中の結 成階間のみからなお呼信器準別を示す。これは、 文中に存在することの目が延点の対象となるもの であって、例えば文中に、恋型切な技練調・が起 込されている場合に掲まする。

第4回は、様文上の評価項目例を示す。左右に 示す文単位の評価項目のからのについて説明する。

- - Φ 衝受け剥り:・格・が正しくない文。例え

ば、モ・悟を、。に、悟にするなどの使い方の文

② 油用中止性:動同の速用形で文が一直切れている場合。例えば、東の単は茎く、大きい、は、鼻(\* (具が長())と、大きい、(泉が大きい)とは、天の寒味が切れている(象の晶が長くて大きい。同様)。

- ② 複雑な逆茂末文 : 長い南置作跡館を持つ文 (複雑な逆茂末文という)。
- の 複雑な複文:従民部を複数持つ文。
- の 二重を定:
- ② 受動態:
- ※ 流通数:1支中に含まれる透陽数。第3回
  (イ)に示すように、少なくても多くても、割り

また、系 4 図中央の間の指摘 X , から X ; , は 1 文中に存在する評価項目のからゆの個数を表し 、右幅の評点 V , から V , , はこれら指揮 X , から

### X,。に対応する呼点を表す。

第5回は、辞在室出电明限を示す。これは、原 (図附位項目のからのに対立する評価器を切から のに対し、加重性のを図示「3"、"2"、" 1"はよび間示評価差別をそ 化変する。加重比 ポロ大きいものほど、博文上の画質に影響を与人 おぼらが大きいことを表している。評価度数は、 評価項目のからの、のからに対し、第3回( の)対点期間(3)のからなとものを通用する。 評価項目のに対し、第3回(4)を形列ののがな いらのを適用する。評価項目のに対し、第3回 (4)を適用する。以下昇点の水出について限率

第5回において、第1に、各将領項目のからの に対応する様文解析データをも、文単位に計判す も、第2に、この計画結果に対し、第5回評領な 地を透用して、対点や、からやい。そ3や計算する。 第3に、計画した辞成や、からやい。これして加ま 比重す。 (n-1 かう10) そ3・東落し、これ ら変質した対象の原料的を評価点として取める。 第4に、この求めた群領点を、第6図右端に示す ように関策して印刷すると共に、必要に応じて孫 6図中央の際に示すように指針文を収集して印刷 せる。

製を図は、本発勢による排文情報、世殊所収お よび指針文例を示す。これは、技術説明器である "FENICSによる解釈VAN頑文"に対し、 本実施例を退期して実際に印刷したリスト例を示 す。関中左から文書号、ページ、行、原文、作針 文、および評価点を示す。磁外に通用した構文ル −ル①ないし⊗のうちの捨立するものは、邪4囚 記載の評価項目①ないし母に対応する。ここで、 技士ルールのは、文の体理が一定の規格に従って いるか否かを評価項目 (株文ルール) にしたもの であって、例えば。住の体質規格。、。領等の体 銀規語。、項の体盤規格。、「例の体数規格。な どが通当であるか否かを評価するものである。ま た、指針文は、第5箇評価項目(精文ルール)の ないし個などに対して算出した個々のほ点が、ほ くなり、雑点質師に入った場合などに、これらに

# 特開平1-220063 (5)

対応するメッセージとして取削したものである。 内、主実は例は、の意定した全ての評価項目を 対象とすることにより、限文上の品質の格価。 なくすることができる。の料価項目のちゃの評 点に対して重みづけ全することにより、西級的な 形況による文書の目文上の品質との差質を少なく することができる。の料価項目の計事報題と、そ 和以外の確点配慮とそ、包含あるいはも影の一般 を利用した資金方式を表入することにより、日本 が取れていまった。

次に、第7回から第10回を用いて、本交明に 係わる文書の意味解釈上の品質の評価について非 細に説明する。

東1回において、日本文作成デバイス 1、文書 ライブラリ 2 は、第1回構成と同じであるので、 お明を古典する。

意味解釈データ収集部15は、日本経解新部1 4によって日本四种番4-1、日本語文法4-2 を参照して解析された対象のうち、文書の重像上 の品質で評価するために必理な意味解収略ポデー タ16年収集するものである。この意味解収略ポデー データ16は、図示のように、不誠一な文体に関 する時間、工具化文学及現に関する情報などがあ

意味品質評価部18は、意味評点重出部11年 よって、意味解析は最テータ18に基づい出資を つけして重出された評点のは初を求めても力する と共に、これらネキの評点(第3間かよび第18 図を用いて後述する)に対応する指針文を出力する ものである。

品質評価リスト19は、室味品質評価部18に よって変換された評価品を、例えば第5回的場合 評価表 192 として示すように、1文単位にリ ストとして印刷したものである。 内、この際、評 個点と耐せて、指針文を第5回や央に示すように メッセージとして回開する。

次に、第8回フローチャートに記述した順序に ほい、第8回、第9回、および第6回を開いて、

#### 城 7 前接皮の動作を評価に抵明する。

第8回において、関中のからのは、第2回回中 のからのに対応するので、説明を省略する。

図の会は、大郎解析、祖文解析、忠学解析、忠 よび選集解析関係を改善する。これは、大郎 大郎 大郎 ない選集解析を行い、文の意思との品質を 詳値するために必要な意味解取時限アータ15年 収集することを意味している。別えば文体、不成 支援、明心表現、共起を現る、別 600ま強(学楽園版、技術、現場、機等、機等 月間域、版等表別、防逆表別、接 時間は、原本表別、防逆表別、降 時間は、原本表別、防逆表別、降 時間は、原本表別、 が成本別などの意味解析場界アー 15を保護する。

図中のは、3・の評価項目ごとの呼流を計算する。これは、"大の憲法"における評価項目の計 調は風に対して、3・の評価項目の許可関係を表 国し、後述する第10回に示すように評点で、な どを計算することを登集している。

関中側は、文の正しさのほ点を行う。これは、

図中ので計画したな。の辞価項目ごとの呼点に対 し、、 ちゃの重みづけした後に、 これらの値を総 計した評価点を算出することを意味している。 個中のは、軍出誌英を編集し、 編集を作成する。

これは、体与れた核果を回復し、高度資産リスト 194日間、耐えば果ら間に示すように日間する ことを理せしている。第5回は、文章はの構成制 を示す。 お娘の神道点質に呼ば点がの前され、中 表に指対ながメッセージとして印刷されている。 以上の処理によって、入力された文章の日本様 が折が行われ、大の意味が収せ減データ16に基 プルで、文書の意味上の高質を定義性がつる とにより、大書の意味上の高質を定義性がつる とにより、大書の意味上の高質を定義的かつ事 他のに背面することが可能となる。

派を図は、定集上の評価項目領を示す。 左関に 采す文単位の評価項目のからのについて説明する。 ① 文体の頃 : " 主えがち"は" です、ま で の吸体。" 本文"は"である側"の等体。 の で表表更 : 別: 従来<u>かか</u> (収集には"か か"の面がままれている)。 約1 8 0 mm <u>でない</u>。

- ② 呼応表現:否定の呼応や仮定の呼応を正しく使用しているか(例:まったく~ない)。
- ② 共認显現:名詞に対する遠疑の選択が正しいか(関:○私は人に会う)。
- Ø 180 · €08:
- の 周辺の限り:用語が誤っていないか(例: -車に)。这話、複合語の近り仮名が正しいか。
- ② 不適切な用値:常用後半の箱間内か。支持を使わないようにしているか (例:しかるに~)。 一般的でない用語を使っていなか (NTから吸い上げる、アミZAP をかける)。
- の 形式名詞: "こと、もの、はず、ため"な どの際し示すものは明確かどうか警告する。
- ② 指示語: "こ、そ、も"が付く指示代名詞の指すものは明確かどうか警告する。
- 毎 用語・その他: "熱調・動詞"の型の複合 当の送り仮名割りなど。
- また、第9回中央の間の指復X,からX。は1 文中に存在する評価項目のから®の開設を表し 、右側の評点V,からV。はこれら指揮X,から

高の間において、個外に適用した意味ルールの ないしゅのうちのほうするものは、来り回記機の が値列目のないしのに対応する。また、設計文は、 明10回評価項目(重味ルール)のないしのなど に対して実出した値々の呼点が、低くなり、様式 別間に入った場合などに、これらに対応する!» セージとして日間引したものである。

#### (発明の効果)

以上既明したように、本効明によれば、入力された文質特に技済文章について様文上および支徴上の品質を加重つきの呼应として選出して品質システムを用いて文書の博文上の品質かよび意味上の品質を定態的に評価することができる。これにより、文書の様文上および意味上の品質を、他の文書と本数的に比較することが可能となると共に、文書の様文上および事業上の品質の均質化を図ることが可能となる。

X。に対応する評点を表す。

2010回は、野点在出設明図を示す。これは、 第3座評価項目のからのに対応する評価店舗のか らのに対し、加重比率を図示 " 3 "、 " 2 "、 \*1 \*および図示評価店単を各・設定する。加数 比平は大きいものほど、意味上の品質に影響を与 える方会が大きいことを表している。評価基準は、 表 3 図 (ロ) 減点疑問(B) のみからなるものを選 用する。以下はみの女用について無具に段明する。 第10因において、第1に、各評価項目のから のに対応する意味解釈データ16を、文単位に計 出する。年2に、この計劃結果に対し、第10回 辞儀器準を適用して、辞点V、からV。を名を計 草する。あるに、計算した経点V、からV。に対 して加速比量で。(n=1から8)を各々東洋し、 これら乗算した結果の無計和を評価点として求め る。気もに、この求めた評価点を、第6回お哨に 示すように編集して印助すると共に、必要に応じ て第6四中央の棚に示すように指針文を編集して

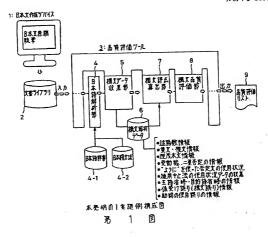
## 4. 西間の簡単な気明

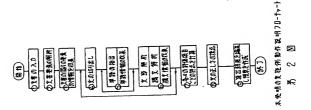
8 M + 4.

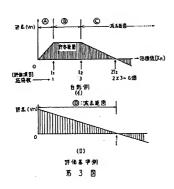
第1回は土泉明の「天津明相級団、東2回は土 発明の支達例制作以明フローチャート、第3回は 対価道路側、第4回は様文上の神師項目側、第3 短は対点単出版明回、第6回は土泉明とよる様文 行標、単電解収などにおける版成例および行針文 削、第7回は土泉明の他の実施例構成回、第8回 は土泉明の他の実施例構成回、平8回 は土泉明の他の実施例構成回、平10回に非依丘 変加減明を示す。

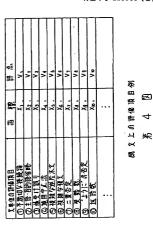
図中、2 は文書ライブラリ、4、1 4 は日本経 訴訴部。5 は様文デーク収集地、6 は様文展学 タ、7 は坊文房反変出版、8 は様文展質降価部 9、19は監督解価リスト、15 は意味解釈結果 デーク収集的、16 は意味解釈結果データ、17 は重年辞点変出版、18 は意味展開解価部を来す。

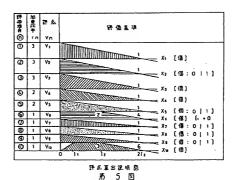
代理人身理士 开始 第一 (外) 年

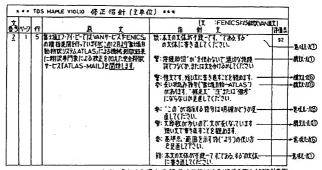






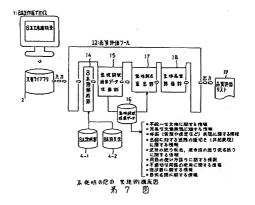


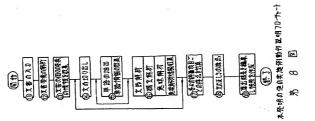




本発明による構文情報、意味解釈などにおける株品のおよび指針文例

**第 6 图** 





文単位の評価項目	指標	7 .S.	
① X (X Ø M) -	Xı	Vı	
文学位の評価項目 ① 文体の統一 ② 亢長表現	X2	V <sub>2</sub>	
③ 呼免表現	Xs	V3	
② 共起表現	X4	V4	
:		T :	

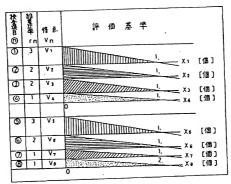
(1)

用語の食。味上の評価項目	19 棟	群 京。	
(5) FA 177 (7) TAY (1)	X 5	V\$	
60不通切な用語	Xé	V6	
分形式为 前	X1	V7	
图 指示器	Xa	Vo	
		:	

(0)

危味上の評価項目例

**第 9 図** 



評価点算出說明图 第 10 图

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS	
$\square$ image cut off at top, bottom or sides	
☐ FADED TEXT OR DRAWING	
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING	
☐ SKEWED/SLANTED IMAGES	
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS	
☐ GRAY SCALE DOCUMENTS	
☐ LINES OR MARKS ON ORIGINAL DOCUMENT	
$\square$ reference(s) or exhibit(s) submitted are poor quality	
D.	

# IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.